THE CCTF WORKING GROUP ON THE EXPRESSION OF UNCERTAINTIES IN PRIMARY FREQUENCY STANDARDS

R.J. Douglas[1] and C. Thomas[2]
on behalf of the Working Group members

[1]National Research Council of Canada, Ottawa, Ontario, K1A OR6, Canada
[2]Bureau International des Poids et Mesures, Pavillon de Breteuil, 92312 Sèvres Cedex, France

Abstract

The Comité Consultatif pour la Définition de la Seconde (CCDS) created in March 1996 a Working Group on the expression of uncertainties in primary frequency standards. This paper presents the main topics included in the first report of the Working Group dated April 1997. This report gives a brief review of the conditions which led to the creation of the Working Group and a summary of the discussions which took place at its first meeting held in Neuchâtel, Switzerland, on 5 March 1997, and also during the following weeks. A main objective of the Working Group is a better understanding between laboratories which evaluate the accuracy of their primary frequency standards and their clients represented by the Bureau International des Poids et Mesures (BIPM) which uses the measurements provided by these standards for assessing the accuracy of TAI (Temps Atomique International, or International Atomic Time) and UTC. An important point for this purpose is the application to primary frequency standards of the guidelines expressed in the* Guide to the Expression of Uncertainty in Measurement *[1].*

CREATION OF THE WORKING GROUP

A primary frequency standard is an instrument which produces an output whose average frequency is based on the internationally accepted definition of the second, and is operated with its own independent implementation of the cesium second and a specified accuracy that does not rely on calibration to another frequency standard. Some national metrology laboratories design and operate large primary frequency standards whose main design objective is to facilitate the evaluation of their accuracy. These instruments are the most accurate devices ever made by humans. The question of quantifying and expressing the accuracy of primary frequency standards - that is "the expression of the uncertainty of primary frequency standards" - arose at the 13th meeting of the CCDS* in March 1996 [2]:

> [Dr Quinn] observed that the *ISO Guide* to the expression of uncertainty in measurements recommends that uncertainties be expressed as a combination of statistical (type A) and others (type B) uncertainties. In view of this, he assumes that the uncertainties quoted for primary frequency standards are entirely type B, since no associated averaging times are quoted.

* *At its 1997 meeting, the Comité International des Poids et Mesures (CIPM) decided to change the name of the Comité Consultatif pour la Définition de la Seconde (CCDS) to that of Comité Consultatif du Temps et des Fréquences (CCTF).*

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **DEC 1997**	2. REPORT TYPE	3. DATES COVERED **00-00-1997 to 00-00-1997**
4. TITLE AND SUBTITLE **The CCTF Working Group on the Expression of Uncertainties in Primary Frequency Standards**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **National Research Council of Canada,Ottawa, Ontario, K1A OR6,**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADA418244. 29th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Long Beach, CA, 2-4 Dec 1997

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**11**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

An extended discussion on the topic followed, which was summarized by Mr Allan who said that in time and frequency applications we usually analyse data in the form of a time series. This leads to the question of what is intended when an uncertainty is calculated, that is, is it the expected variation of a quantity over a period of time, or is it merely the uncertainty of a single measurement?...

Dr Winkler said that the evaluation of a frequency standard involves two steps: the determination of systematic effects and then the measurement of its frequency with respect to TAI. This second step involves the consideration of measurement times and instability. Prof. De Marchi agreed with Dr Winkler, saying that stability and accuracy are very different entities...

Dr Winkler gave the opinion that the *ISO Guide* is applicable so long as the distinction is made between stability and accuracy.

Dr Sullivan added that some of the effects encountered in the uncertainty budget of a primary standard may not be independent. In such a case the quadratic sum of the uncertainty components is not sufficient.

Dr Bauch added that it is sometimes difficult to distinguish between type A and type B uncertainties; for example, the uncertainty in the first order Doppler effect in the fountain frequency is evaluated by measurement and thus cannot be treated as a type B uncertainty.

... the President... suggested that the matter is of sufficient importance that the CCDS should form a Working Group to report on how the accuracy of primary frequency standards should be evaluated in accordance with the *ISO Guide*. He asked Dr Douglas to act as Chairman of the CCDS Working Group on the expression of uncertainties in primary frequency standards... the membership of the Working Group is... open to... experts from timing laboratories.

The first meeting of this Working Group was held on 5 March 1997 at the 11th European Forum on Time and Frequency, in Neuchâtel, Switzerland. A preliminary exchange had been carried out via e-mail between those attending the first meeting and several other experts.

SCOPE OF THE WORKING GROUP

The scope of the Working Group was approached by its Chair, in an e-mail dated 22 January 1997, addressed to the Group members:

...it has become evident to me that it is vital to choose the correct scope: one that could result in our recommending a way to communicate the basis for frequency metrology rigorously and widely, without overburdening or bypassing any particular part of the frequency metrology chain - from primary standard metrologist to end-user.

In this approach, it is emphasized that user needs for frequency sources, though not necessarily at the upper level of accuracy, are important and should be discussed by the Working Group. It was suggested that the word 'primary' be dropped from the group title.

However, the discussions inside and after the first meeting led to the conclusion that the first objective of the Working Group is to develop a better understanding between those who are building and evaluating primary frequency standards, and those who are using them, one of the main users being the BIPM. Direct users, and especially the BIPM would be more fully informed on the evaluation of primary frequency standards, to allow the standards' data to be used in an optimum way. The Working Group discussed extending the scope to include the use of the results of primary frequency standards for calibrating instruments (with reference to TAI or UTC, for example). It was decided to restrict the scope of the Working Group's report for the next CCTF meeting to the expression of uncertainty of primary frequency standards, particularly uncertainty budgets aimed at frequency transfer to TAI.

TYPE A AND TYPE B UNCERTAINTIES

There exists a formal and published recommendation, Recommendation INC-1 (adopted in 1980) on the statement of uncertainties. This recommendation, on the expression of experimental uncertainties is the basis of the *ISO Guide to the Expression of Uncertainty in Measurement* [item 0.7, p viii]:

> 1. The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:
>
> A. those which are evaluated by statistical methods,
>
> B. those which are evaluated by other means.
>
> There is not always a simple correspondence between the classification into categories A or B and the previously used classification into 'random' and 'systematic' uncertainties. The term 'systematic uncertainty' can be misleading and should be avoided.
>
> Any detailed report of the uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical value.
>
> 2. The components in category A are characterized by the estimated variances s_i^2 (or the estimated 'standard deviation' s_i) and the number of degrees of freedom ν_i. Where appropriate, the covariances should be given.
>
> 3. The components in category B should be characterized by quantities u_j^2, which may be considered as approximations to the corresponding variances, the existence of which is assumed. The quantities u_j^2 may be treated like variances and the quantities u_j like standard deviations. Where appropriate, the covariances should be treated in a similar way.
>
> 4. The combined uncertainty should be characterized by the numerical value obtained by applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of 'standard deviations' (referred to as *standard uncertainties* since the publication of the *ISO Guide* in 1993[1]).
>
> 5. If, for particular applications, it is necessary to multiply the combined uncertainty by a factor to obtain an overall uncertainty, the multiplying factor used must always be stated.

The division of evaluation methods for uncertainties into type A and type B methods is in part an echo of the "random" and "systematic" uncertainties of past decades. However, since there is no difference in the way the two types are to be used, the type A and type B classification is only an issue for communicating the methodology rather than the value of the uncertainty.

STABILITY OF PRIMARY FREQUENCY STANDARDS

Most of the *ISO Guide* takes the implicit condition that the system under study is stationary and affected only by white noise. This is obviously not the case for time measurements which often deliver time series affected by correlated noise such as random walk of frequency. Expressing a type A uncertainty in the form of a generic standard deviation s, as recommended by the Guide, makes no sense in this case. To construct a stationary standard uncertainty, a specific frequency holdover pattern can be considered, calibrated over one time interval τ_1 and used over another time interval τ_2 a time t later [3,4]. Used in this way, tools like the Allan variance, $\sigma_y^2(\tau)$[5], are suitable for specifying the consequences of known non-white noise on the uncertainty of primary frequency standards where deterministic frequency drift is not a major effect. When it is important, deterministic drift can be handled in similar ways [6]. The question of the Allan variance in time measurements is evoked in item 4.2.7 of the *ISO Guide*:

> If the random variations in the observations of an input quantity are correlated, for example, in time, the mean and experimental standard deviation of the mean ... may be inappropriate estimators of the desired statistics. In such cases, the observations should be analysed by statistical methods specially designed to treat a series of correlated, randomly-varying measurements.

NOTE - Such specialized methods are used to treat measurements of frequency standards... (See reference [5], for example, for a detailed discussion of the Allan .variance).

Uncertainty related to stability is thus treated with type A methods, but it may be possible to classify other sources of uncertainty in primary frequency standards as being also of type A.

Applied to some evaluation methods used in modern primary frequency standards, the type A and type B classification is ambiguous. Working Group members report different interpretations which result in strikingly similar methods, used by different laboratories, being classified by some as of type A, and by others as of type B.

For a primary frequency standard, all perturbing effects are best determined by measurement. The measurements are repeated, and evaluated - usually with the aid of a theoretical model whose parameters are determined by fitting to measured values - and the correction to zero perturbation is determined from the fit. The fitting process is certainly a statistical method, and viewed in this way the derived uncertainty in the extrapolated correction can then be described as having been evaluated by a type A method. However, the use of a theoretical model, and perhaps ancillary apparatus for determining the correction, can lead to the view that the correction has been imported and not determined by statistical analysis of the primary measurand - and so it has been evaluated by a type B method.

This classification, meant to communicate information about the evaluation method, is actually obscuring the similarities of the methods used. Thus the Working Group will have to address the refinement of this classification system, or consider abandoning it altogether. Abandoning the classification has implications for other fields where statistical methods and other knowledge interact. Even the purest type A method imports the model of a stationary mean, so it too might be viewed as using in part an "other method". Of those who were helped develop this classification into type A and type B evaluation methods, there is a willingness to consider abandoning the distinction [7]. The essential element is that uncertainty components, evaluated by any method, are to be construed as probability distributions and combined in the same way. The type A and type B classification filled the important role of replacing the "random" and "systematic" nomenclature, which were not always treated together, and unifying the treatment of their uncertainty distributions in a single uncertainty. With this primary function largely completed, it may be time to consider de-emphasizing the classification.

AVERAGING TIME FOR PRIMARY STANDARD EVALUATION

An important point about the question of frequency measurements and the evaluation of the uncertainty budget for a primary frequency standard concerns the averaging time chosen for the operations of evaluating and reporting average frequencies. The situation is not the same for the diverse very accurate primary frequency standards in operation in the world:

- The averaging time chosen by the American national metrology laboratory, the National Institute for Standards and Technology (NIST) for its optically pumped thermal beam primary frequency standard, NIST-7, is basically set by the sampling of TAI, 10 days in 1994 and 1995 and 5 or 10 days since January 1996. This makes it possible to cancel the frequency variations of the local hydrogen maser to which NIST-7 is compared, and thus ensures an optimum link to TAI.
- The German national metrology laboratory, Physikalisch-Technische Bundesanstalt (PTB), operates thermal beam primary frequency standards which use magnetic state selectors in an axial field geometry. Their standards are operated continuously and their data are sent to the BIPM as for conventional clocks contributing to TAI. The choice of

two-month periods for the estimation of their frequencies relative to TAI was made by the BIPM, simply because TAI is computed with blocks of two-month data.

- The French national metrology organization, the Bureau National de Métrologie (BNM), through its Laboratoire Primaire du Temps et des Fréquences (LPTF), operates a cesium fountain primary frequency standard. It is not operated continuously. The averaging time chosen for the BNM-LPTF cesium fountain FO1 is not related to the sampling time of TAI. The level of white frequency noise in FO1 is so low that, for an averaging time of about 10 hours, its curve of stability is limited by the stability of the local hydrogen maser to which FO1 is compared. Complete evaluation of the standard cannot be carried out efficiently over longer averaging times.

It is often supposed that the accuracy of a primary frequency standard is evaluated in the range of averaging times corresponding to its best stability, but it is not always the case:

- It is not possible for FO1 since the stability of the local oscillator limits the measurements.
- For what concerns NIST-7, a complete evaluation is carried out each time a new frequency measurement is provided to the BIPM, and is thus valid over averaging times of 5 or 10 days. The part of the type A uncertainty coming from the lack of stability of the standard is usually small compared to some of the uncertainties associated with the different frequency corrections, but that this is not always or necessarily true.
- There is no search to make regular complete evaluations of the accuracy of the PTB standards each two months: this is done at the beginning and is well documented [8]. Some contributions are routinely and frequently re-evaluated (mean magnetic field, second order Doppler, electronics, microwave spectrum) and the claimed uncertainties are repeatedly verified. Others are estimated only once (magnetic field inhomogeneity, cavity pulling, Rabi pulling, gravity, black-body). For what concerns the end-to-end cavity phase shift, it is checked from beam reversals; this specific uncertainty, of type B, is based on mechanical measurements made during construction of the clock, which cannot be repeated unless the clock is dismantled.

TRANSFER OF FREQUENCY MEASUREMENTS TO TAI

One important use of primary frequency standard measurements is for the assessment of the accuracy of TAI. It is thus essential to optimize the internal link between the standard and the time-transfer measurement system, often a GPS time receiver, inside the laboratory, and to treat with precautions the time transfer data to TAI:

- Links are optimized for NIST-7 thanks to the choice of averaging periods which match the TAI time grid. However, the component of uncertainty due to the GPS transfer over 5 days may not be negligible when compared with the standard uncertainty of NIST-7 as given at the end of 1996 (5×10^{-15}).
- The standard PTB CS2 provides UTC(PTB) and is used directly as input to the PTB GPS receiver. Since data are averaged over two-month periods, TAI takes advantage of the full accuracy of the PTB standards.
- After transfer to TAI, the FO1 data of September-December 1995 show a huge dispersion of the measurements (peak-to-peak 2×10^{-14}) when compared to the total uncertainty of the standard at that time (3×10^{-15}). The internal laboratory link was not optimized at that period, but this should be done for future measurements.

A primary frequency standard measurement is carried out over a given time interval, defined by its duration and its central date. The measurement calibrates the rate of a local time scale, which is compared (largely by GPS common-view techniques) with the time scales of other laboratories in the pooling of comparisons for the determination of the paper time scale TAI. The measurement can thus be used to calibrate the rate of TAI over a time interval with

another duration and another central date. The frequency holdover is carried out using the stability of TAI itself: the frequency measurement is transferred without change, but an additional uncertainty term is included (added in quadrature using the root-sum-square) to account for the relative stability of the two time scales, scaled by an amount depending upon the noise characteristic of the time scale difference [3,4].

UNCERTAINTY BUDGETS FOR PRIMARY STANDARDS

It is agreed that the uncertainty of a primary frequency standard should be documented in tabular form. This table should list all corrections applied to the standard data, together with their respective uncertainties, and references to the procedures used for measuring influence parameters and references to the functional forms and numerical constants used for calculating corrections. These references are important for allowing the evaluation of correlations between the methods used in different primary standards. Where appropriate the degrees of freedom should also be specified, to indicate how well the uncertainty has been determined. Together, this information constitutes the 'uncertainty budget'. This procedure is in complete accord with the *ISO Guide*. The elements of the uncertainty budget for a primary frequency standard have more in common with a pure physics measurement of a fundamental constant (the hyperfine interval of ^{133}Cs) than with classical metrology. They are covered in extensive textbooks on the subject [9], and is an active area of continuing research. They are touched on below. The strength of the *ISO Guide* is evident in the ease and completeness with which it encompasses the expression of uncertainty for both fundamental constants and classical metrology. Much of the Working Group's discussions will use the *ISO Guide* as their primary reference.

On several occasions the BIPM has published in its regular *Circular T*, a value of the duration of the TAI scale unit obtained from a primary frequency standard measurement, with the uncertainty budget of the standard, before the corresponding evaluation of accuracy could be published or even submitted to an international journal. The BIPM thus serves as publisher for results which may be revealed later to be too optimistic or too quickly obtained. The Working Group may decide that the BIPM should always indicate the source of information containing the claimed uncertainty, or re-publish uncertainty budgets that it uses.

Returning to the general problem of drawing up the uncertainty budget for a primary frequency standard, one may wonder if there is a complete consensus about what corrections should be listed in the table. One example that illustrates how there can be a difference in practice, rooted in a philosophical difference, is the gravitational shift.

One view is that since a frequency standard produces a proper unit, the SI second, defined locally from an experiment within the dimensions of a laboratory, the gravitational shift is small enough so that it should not be included in the uncertainty budget. For the largest cesium fountains ever seriously considered [10,11], the effect within the standard has a maximum difference of 0.5×10^{-15}, while for horizontal thermal beam cesium standards the maximum difference may be smaller than 1×10^{-18}. The uncertainty in the difference in gravitational potential within the laboratory will normally beeasilyat a negligible level.

A second view is that in addition to these small corrections and even smaller uncertainties, the correction and its uncertainty should be taken into account when the SI second is transferred to the rotating geoid for assessing the accuracy of TAI. For clients of primary frequency standards to be served in this way, a laboratory takes the responsibility of estimating this part of the uncertainty (which they are best equipped to do) for the way their frequency standard is used outside the confines of their laboratory. Proper time users also

have a gravitational correction to estimate. These corrections [12] can be large - about 10^{-9} for some Earth orbits. Uncertainties can range from 1×10^{-18} in Earth orbit to 1×10^{-17} for carefully surveyed elevations on the Earth's surface, to 10^{-15} if one uses a topographic map or a simple ellipsoid to geoid conversion, and can approach 10^{-14} if elevation above the ellipsoid is used in place of elevation above the geoid (i.e. elevation above sea level).

Different laboratories do not treat this question in the same way: the gravitational correction appears as an additional line in the FO1 uncertainty budget, after the quadratic sum of the different uncertainties has been made; for other standards it is taken into account in the main table. A consensus on this question has not developed in the Working Group.

For some members, it is understood that if people are developing standards it is largely for use at other locations in the outside world, so a reference location is chosen to convert the produced SI second, a unit of proper time, to a unit of coordinate time. The rotating geoid is almost universally chosen for terrestrial applications, most often through TAI or UTC. In this perspective the laboratory estimates and applies the gravitational correction and its uncertainty as an intrinsic part of realizing the unit of TAI.

Other members recall that the second is defined as a proper unit, and would classify the gravitational uncertainty in the transfer to TAI as an issue removed from the generation of the scale unit of a proper time scale. Some have suggested that the uncertainty in the gravitational correction is so small that this problem is not yet important.

In principle there should be no problem in agreeing about the corrections to be considered in the uncertainty budget. However, it should be recalled that the correction for the black-body shift had been neglected for years: it has been applied uniformly since 1995, although the paper first describing the effect was published in 1982 [13]. The first formal action on the effect was by the 1985 meeting of the CCDS, which recalled that the definition of the second was for the limit of zero perturbations and called for studies on these matters [14]. Although the theoretical underpinning seemed solid enough, experimental verification that the effect scaled as expected with blackbody temperature was not available until 1996 [15,16]. This led to a situation where there were two approaches in use: one held that the theory (and the old low frequency Stark shift measurements on cesium [17,18]) was the best estimate of the blackbody radiation shift, and the other held that further experimental confirmation on cesium atoms with appropriate blackbody radiation was necessary before the correction should be applied. The typical shift (at 300 K about -2×10^{-14}) is small enough to be difficult to measure and yet large enough to become a major source of uncertainty if no correction is applied. The uncertainty budgets of primary frequency standards through the decade 1985-1995 were not always published frequently enough nor widely enough to make it easy to understand how the blackbody radiation was being treated for widely accessible standards.

About the correction for the black-body shift, discussions took place inside the first meeting of the Working Group on whether its uncertainty should be classified as type A or as type B. The evaluation is based on the formula given by Itano *et al* [13]:

$$[\nu(T)-\nu_0]\,/\,\nu_0 = -\,(169\pm4)\times10^{-16}\,(T/300)^4.$$

with $\nu(T)$ the frequency at temperature T (T in K) and ν_0 the frequency at 0 K. At room temperature, the correction is -1.7×10^{-14} with a variation of -2.3×10^{-16} per K.

This shift has not yet been measured directly for the temperature of primary standards as compared to the unperturbed low temperature limit (T=0 K), and we do not have yet a complete understanding of the limitations of the formula. The uncertainty quoted above

arises partly from previous measurements of the scalar low frequency Stark effect but also makes some allowance for inadequacies of the theoretical treatment. Many kinds of expertise are needed to estimate the extent to which these are the only sources of uncertainty. The formula's uncertainty should incorporate tensor effects as well as the AC Zeeman shift effect. Estimating the effective temperature of the radiation calls for knowledge of the temperature of the surroundings, and if the temperature distribution is not uniform, a knowledge of the emissivity of the surfaces. The uncertainty of the correction for the black-body shift is generally estimated to be 1×10^{15} for standards operating at 300 K. This is a conservative value relative to the uncertainties given in the formula and in the temperature, and should thus be regarded as of type B. Such a discussion should be undertaken for all of the uncertainty components which appear in the uncertainty budget of primary standards.

The largest correction term, which also has an associated uncertainty, is the Zeeman shift deliberately induced by a small magnetic field applied to the cesium atoms. The field is applied to isolate and use the most stable component of the cesium hyperfine transition (between $F=3$, $m_F=0$ and $F=4$, $m_F=0$), which has a quadratic field dependence and a shift that is typically 2×10^{-9} or less. Most often the magnetic field is measured with the neighboring components of the cesium hyperfine transition (such as between $F=3$, $m_F=0$ and $F=4$, $m_F=1$). Evaluating and describing the stability and homogeneity of the magnetic field, and the uncertainty of the correction which these variations generate, will not be done identically for different types of standard. The completeness of the Breit-Rabi formula [9] for these transitions is an interesting issue, since any revision here could affect, unequally, standards which use different fields, and yet if the corrections are properly documented today's biases from this type of uncertainty could be removed retrospectively in later redeterminations of TT for astronomical purposes.

The Stark shift is generally negligibly small, since to keep this correction less than 10^{-16}, the electric field only needs to be kept below 5 V/m.

The second-order Doppler shift or time dilation is the second largest correction to be applied for many thermal beam frequency standards, since for thermal cesium atoms at 300 m/s the effect is $v^2/c^2 \approx 10^{-12}$. Determining the speed distribution of the cesium atoms can be done in different ways, and can be done with an uncertainty of 10^{-15} for thermal beams. For laser-cooled cesium beams the correction is normally less than 10^{-15} and the speed distribution can usually be measured even more accurately so that this uncertainty could be 10^{-18} or less.

Collisions in the vacuum in which the cesium atoms drift can affect the resonance frequency. The largest, and to some extent unavoidable, effect arises from collisions of the beam with other cesium atoms in the beam which gives a density shift with an uncertainty which might be evaluated with an uncertainty of 10^{-16} for cesium fountains.

Distributed phase variations in the microwave cavity give a phase shift for the two passages of the atoms through the cavity, with care room temperature cavities can be evaluated to produce an uncertainty of 10^{-6} of the linewidth or less. The time-averaged Poynting vector can be reduced by selecting cavity design and wall materials. The critical element is the absence of net energy flow rather than the Q of the cavity. Ring cavities, symmetric feeds, and in the future perhaps superconducting cavities may be used .

Other transitions than the intended one (between $F=3$, $m_F=0$ and $F=4$, $m_F=0$), can also affect the uncertainty. The cesium hyperfine transition is not a two-level one, but there are seven Zeeman sublevels of F=3 and nine sublevels of the F=4 level. The sublevels are purposefully split by the applied magnetic field, and excitation of unwanted transitions can lead to Rabi pulling or Ramsey pulling of the apparent center of the intended transition. Evaluating the

unwanted transitions involves considering the purity of the initial state preparation, the parallelism of the applied static and microwave magnetic fields, the spectral purity of the microwaves and the ability of the final state analyzer to reveal the final superposition of states. Describing the procedures used to evaluate all this can clearly be a challenge.

Microwave cavity leakage can affect the standard, and the evaluation of any residual microwave fields in the drift space leads to another source of uncertainty.

The main frequency control servo must be carefully characterized in terms of its offset. Unwanted effects arise from any unwanted detector signal which is coherent with the frequency modulation or switching of the servo system, but are easy to characterize. For analog servos the inverter, chopper and zero offset of the main integrator must be examined particularly carefully. The effects of servo design, bandwidth and gain are also considered, and can be modelled with the local oscillator to give a predicted stability which can be considered as a part of the uncertainty budget, particularly when describing the variation, with averaging time, of the standard uncertainty of the primary frequency standard.

The local oscillator's phase and frequency noise combines with the shot noise of the interrogation of the cesium resonance to affect the stability component of the standard uncertainty of the primary frequency standard, changing the servo response and (usually) to a lesser extent changing the excitation of unwanted transitions and uncertainty budget.

Environmental effects alter the physical environment of the cesium atoms of a primary frequency standard in measurable ways - measurable through the way that they change the above influence parameters of the cesium atoms. Environmental variations which are unanticipated may not be fully captured and accounted for by scheduled measurements. For a primary frequency standard there will be strategies for identifying random, diurnal, and seasonal effects and re-measuring corrections varied by magnetic fields, RF interference, temperature, temperature gradients, humidity, and atmospheric pressure. These variations will be associated with uncertainty in the influence parameters and so in the final frequency.

The transfer characteristics of the average frequency of a primary frequency standard to TAI is another area for active discussion. The laboratory possessing a primary frequency standard which is to be used in steering TAI is best equipped both to design the transfer strategy and to evaluate the uncertainty associated with the transfer of the average frequency to TAI. The Working Group will likely consider ways of exploiting this expertise for documenting and perhaps improving the evaluation of the frequency of TAI relative to the SI second.

CONCLUSIONS

In the first meeting of the CCTF Working Group on the expression of uncertainties in primary frequency standards, it was clear that those who produce measurements from primary frequency standards and those who use them do not always have a complete appreciation on what has been done in other laboratories. The Working Group could recommend a framework for the publication of uncertainty budgets and other reports on primary frequency standards. Uncertainty budgets of primary frequency standards would be given in accordance with the *ISO Guide*, and with some supplementary material specific to frequency metrology. In this framework several topics remained to be studied, in particular the classification of uncertainties into types A and B, and the procedures by which uncertainty components should be combined and weighted averages taken.

MEMBERS OF THE WORKING GROUP

It is a pleasure to acknowledge those who have participated in the deliberations of the Working Group to date: D. Allan, A. Bauch, J.-S. Boulanger, A. Clairon, A. De Marchi, R. Drullinger, P. Fisk, A. Lepek, and T. Parker.

REFERENCES

[1] *Guide to the Expression of Uncertainty in Measurement*, 101 p (International Organization for Standardization, ISBN 92-67-10188-9, 1993).

[2] Report of the 13th Meeting of the CCDS, *BIPM Publications*, item 2.2 p S56, (1996).

[3] J. Azoubib, M. Granveaud, B. Guinot, *Metrologia*, **13**, 87-93 (1977).

[4] R.J. Douglas, J.-S. Boulanger, *Proc. 11th EFTF*, 345-349 (1997).

[5] D.W. Allan, *IEEE Trans. Instrum. Meas.* **IM 36**, 646-654 (1987).

[6] R.J. Douglas, J.-S. Boulanger and C. Jacques, *Proc. 25th PTTI*, 249-266 (1993).

[7] P. Giacomo, personal communication to the Working Group, (1997).

[8] A. Bauch, T. Heindorff, R. Schröder, B. Fischer, *Metrologia* **33**, 3, 249-259 (1996).

[9] J. Vanier and C. Audouin, *The Quantum Physics of Atomic Frequency Standards* (Adam Hilger, Bristol and Philadelphia, 1987).

[10] J.R. Zacharias, *Phys Rev* **94**, 751 (1954).

[11] A. De Marchi, *Metrologia* **18**, 103-116 (1982).

[12] P. Wolf, *Relativity and the Metrology of Time*, BIPM Monograph 97/1, (1997).

[13] W.M. Itano, L.L. Lewis, D.J. Wineland, *Phys Rev A*, **25**, 1233-1235 (1982).

[14] Report of the 10th Meeting of the CCDS, *BIPM Publications*, S10, S44, and S58 (1985).

[15] A. Bauch and R. Schröder, *Phys Rev Lett* **78**, 662 (1997).

[16] R. Augustin, A. Bauch and R. Schröder, *Proc. 11th EFTF*, 47-52 (1997).

[17] R.D. Haun and J.R. Zacharias, *Phys Rev* **107**, 107 (1957); J.R. Mount, *Phys Rev* A **5**, 1059 (1972).

[18] E. Simon, P. Laurent, C. Mandache and A. Clairon, *Proc. 11th EFTF*, 43-46 (1997).

Questions and Answers

CLAUDINE THOMAS (BIPM): Maybe I can make one comment. Thank you for showing us the guide on the expression of uncertainties. Please do it again.

ROB DOUGLAS (NRC): That is what it looks like in the light.

CLAUDINE THOMAS: It is a very important book for us, of course, because it explains how we must express our uncertainties and combine them, of course; so thank you for that.